

An Integrative Bioinformatics Approach to Study Single Cancer Cell Heterogeneity

The University of Hawaii at Manoa

PI: Lana X. Garmire

Grant Number: 1-K01S25434-01

My long term career goal is to become a leading expert in translational bioinformatics who creates, develops and applies computational and statistical methods to reveal landscapes of cancers and to identify strategies to cure cancers. Human cancers are highly heterogeneous. Such heterogeneity is the major source of the ultimate failure of most cancer agents. However, due to the limit of technologies, the intercellular heterogeneity has not been investigated genome wide, at single-cell level until recently. New technologies such as single-cell transcriptome sequencing (RNA-Seq) and exome have revealed new insights and more profound complexity than was previously thought. However, so far these technologies are limited to one assay per cell. It remains a grand challenge to perform multiple, integrative assays from the same single tumor cell, in particular, from those derived from small tumor biopsies. Given the stochasticity at the single cell resolution, reproducibility and sensitivity are daunting tasks. To overcome this challenge, I have started the single cancer cell sequencing analysis project, in collaboration with Dr. Sherman Weissman at Yale University, who is also my co-mentor of this K01 proposal. My immediate career goal is to identify genome-wide heterogeneity among single cancer cells, using the erythroleukemia K562 cell line. Towards this, I am proposing a research project on an integrative bioinformatics approach to analyze multiple types of genomics data generated from the same single leukemia cells, a timely and critical topic. Specifically, I am interested in studying the following specific aims: (1) building a bioinformatics pipeline to study heterogeneity of single-cell RNA-Seq, (2) building a bioinformatics pipeline to study CpG methylome of single cells, (3) building a bioinformatics pipeline to study single-cell Exome-Seq, and (4) integrate the RNA-Seq, methylome and Exome-Seq data generated from the same single cells. These single cells genomic data are provided by Dr. Sherman Weissman's lab from 30 single K562 erythroleukemia cells. I will first construct and validate in parallel, the RNA-Seq, methylome, and Exome-Seq bioinformatics pipelines optimized for single-cell analysis, and then develop and validate an integrative platform to analyze these multiple types of high-throughput data. To accomplish the research project, and to successfully transit from a junior faculty to an expert of the field, I have developed a career plan with my mentoring committee composed of four world-class experts in different fields relevant to Big Data Science: Primary Mentor Dr. Jason Moore in Bioinformatics from Dartmouth College, Co-mentor Dr. Sherman Weissman in Single-cell Genomics and Genetics from Yale University, Co-mentor Dr. Herbert Yu in Cancer Epidemiology from University of Hawaii Cancer Center and Co-mentor Dr. Jason Leigh in Big Data Visualization from the Information and Computer Science Department of University of Hawaii Manoa. I will primarily work with my four co-mentors for planning the development of my career during this award. PUBLIC HEALTH RELEVANCE: The goal of this K01 proposal is to integrate multiple types of high-throughput data, in particular, the transcriptome, exome-sequencing and CpG methylome data generated from single cancer cells. The proposed project is designed to address the urgent need for an integrative bioinformatics platform for mega-data generated from next-generation sequencing applications. It is also aimed to study the fundamental sources of tumor heterogeneity.